



(11) Publication number: **2002373174 A**

Generated Document

(21) Application number: 2001182122

(51) Intl. Cl.: G06F 17/30 G06F 12/00

(22) Application date: 15.06.01

(30) Priority:

(43) Date of application publication: 26.12.02

(84) Designated contracting states:

(71) Applicant: NEC CORP

(72) Inventor: MIYAZAKI TETSUJI

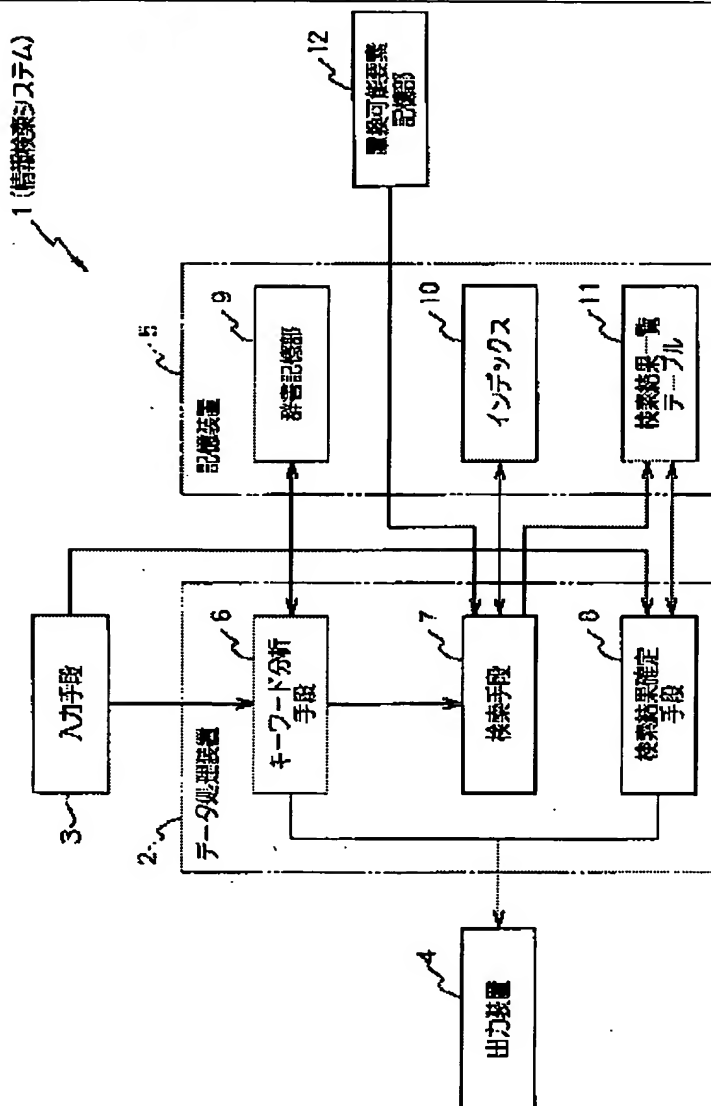
(74) Representative:

(54) SYSTEM AND PROGRAM
FOR RETRIEVING
INFORMATION

(57) Abstract:

PROBLEM TO BE SOLVED: To provide an information retrieving system for allowing even the beginner of the Internet to easily know the location of necessary information in a short time.

SOLUTION: A keyword inputted as a sentence is syntax-analyzed by a keyword analyzing means 6 so that the keyword can be divided into primary constituting elements and the other constituting elements, and keyword retrieval is executed by a retrieving means 7 by using only the primary constituting elements as a keyword in the first stage so that a mother group being the targets of retrieval can be retrieved and temporarily stored in a retrieved result list table 11. Afterwards, the keyword retrieval is executed within the range of the mother group by using the other constituting elements constituting a part of the keyword for narrowing-down the retrieved results and replaceable elements replaceable for them so that any waste that an index 10 with a large capacity is re-retrieved can be eliminated, and that the keyword retrieval using a precise retrieval condition can be executed in a short time.



(11)特許出願公開番号

特開2002-373174

(P2002-373174A)

(43)公開日 平成14年12月26日(2002.12.26)

(51) Int.Cl.⁷

識別記号

FI

テーマコード(参考)

G 0 6 F 17/30

330

G O 6 F 17/30

330C 5B075

12/00

520

12/00

520E 5B082

審査請求 有 ・ 請求項の数 7 O L (全 12 頁)

(21)出願番号 特願2001-182122(P2001-182122)

(22)出願日 平成13年6月15日(2001.6.15)

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 宮崎 哲司

東京都港区芝五丁目7番1号 日本電気株式会社内

(74)代理人 100079164

弁理士 高橋 勇

Fターム(参考) 5B075 NK06 NK32 PQ15 PR08

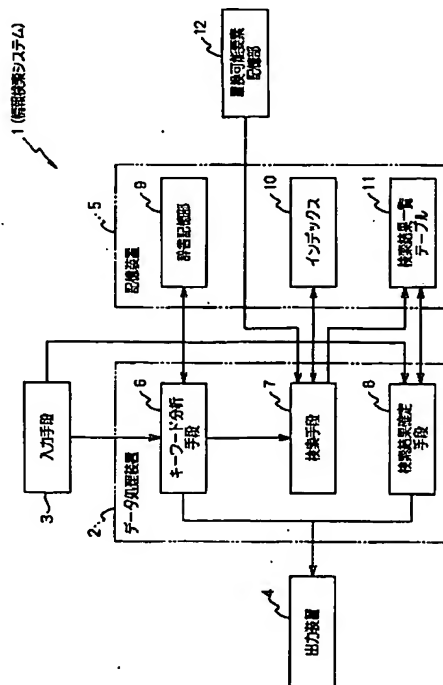
5B082 EA08 GC04

(54) 【発明の名称】 情報検索システムおよび情報検索プログラム

(57) 【要約】

【課題】 インターネットの初心者であっても、必要とする情報のありかを短時間のうちに容易に知ることのできる情報検索システムを提供すること。

【解決手段】 センテンスとして入力されたキーワードをキーワード分析手段6で構文解析することによって主要な構成要素と他の構成要素とに分割し、まず、最初の段階で主要な構成要素のみをキーワードとして検索手段7による検索を実行して検索対象とする母集合を求めて検索結果一覧テーブル11に一時記憶する。その後、検索結果を絞り込むためのキーワードの一部を構成する他の構成要素およびこれに代替可能な置換可能要素を用いて母集合の範囲内で検索を実行することで、大容量のインデックス10が再検索される無駄をなくし、詳細な検索条件を利用したキーワード検索を短時間で実施できるようにする。



【特許請求の範囲】

【請求項1】 情報のありかとキーワードとを対応させて記憶したインデックスを利用し、入力されたキーワードに対応する情報のありかを検索して提示する情報検索システムであって、

センテンスとして入力されたキーワードを構文解析して主要な構成要素と他の構成要素とに分割するキーワード分析手段と、

前記主要な構成要素のみをキーワードとして前記インデックスを検索することにより検索対象の母集合となる検索結果を求めて検索結果一覧テーブルに一時記憶させると共に、前記他の構成要素および該他の構成要素に代えて利用することのできる幾つかの置換可能要素をキーワードとして前記検索結果一覧テーブルを検索して前記検索結果一覧テーブル内の検索結果をグループ化する検索手段と、

前記他の構成要素および前記幾つかの置換可能要素の中から択一的に要素を選択するための入力手段と、

前記入力手段によって選択された要素に対応してグループ化されている検索結果の一覧を提示する検索結果確定手段とを備えたことを特徴とする情報検索システム。

【請求項2】 前記幾つかの置換可能要素を前記他の構成要素に対応させて記憶した置換可能要素記憶手段を前記検索手段に併設したことを特徴とする請求項1記載の情報検索システム。

【請求項3】 前記キーワード分析手段は、名詞および固有名詞を主要な構成要素として認識する主要構成要素認識機能を備えていることを特徴とする請求項1または請求項2記載の情報検索システム。

【請求項4】 前記検索結果確定手段は、前記検索手段の作動終了時に、前記他の構成要素を自動的に前記要素として選択する最適検索結果提示機能を備えていることを特徴とする請求項1、請求項2または請求項3記載の情報検索システム。

【請求項5】 前記検索手段は、分割された前記他の構成要素毎に検索優先度を判定し、検索優先度が高い前記他の構成要素およびこの構成要素に代えて利用することのできる幾つかの置換可能要素のみをキーワードとして前記検索結果一覧テーブルを検索して検索結果一覧テーブル内の検索結果をグループ化する検索条件最適化機能を備えていることを特徴とする請求項1、請求項2、請求項3または請求項4記載の情報検索システム。

【請求項6】 前記入力手段は、前記他の構成要素および前記幾つかの置換可能要素を並列してモニタ上に表示する選択可能要素一括表示機能を備えていることを特徴とする請求項1、請求項2、請求項3、請求項4または請求項5記載の情報検索システム。

【請求項7】 情報のありかとキーワードとを対応させて記憶したインデックスを利用し、入力されたキーワードに対応する情報のありかを検索して提示する情報検索

システムの主要部を構成するコンピュータに設置される情報検索プログラムであって、

前記コンピュータにセンテンスとして入力されたキーワードを構文解析して主要な構成要素と他の構成要素とに分割するキーワード分析機能と、

前記主要な構成要素のみをキーワードとして前記インデックスを検索することにより検索対象の母集合となる検索結果を求めて検索結果一覧テーブルに一時記憶させると共に、前記他の構成要素および該他の構成要素に代えて利用することのできる幾つかの置換可能要素をキーワードとして前記検索結果一覧テーブルを検索して前記検索結果一覧テーブル内の検索結果をグループ化する検索機能と、

入力手段または情報伝達経路を介して入力される指令に応じて前記他の構成要素および前記幾つかの置換可能要素の中から択一的に要素を選択し、該選択した要素に対応してグループ化されている検索結果の一覧を出力する検索結果確定機能とを前記コンピュータによって実現するための情報検索プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、情報検索システムおよび情報検索プログラムの改良、特に、目的とする情報を短時間で容易に取得するための改良に関する。

【0002】

【従来の技術】家庭へのパーソナルコンピュータの普及に伴ってインターネットの利用者の数も順調に伸びてきている。インターネットを利用することにより世界中の情報にアクセスできるようになったが、その反面、本当に必要な情報を探すのも大変になってしまった。

【0003】

【発明が解決しようとする課題】現在の情報検索システムおよび情報検索プログラムであっても、熟練者が使用すれば必要な情報に簡単にアクセスすることができる。しかし、初心者の場合には、必要とする情報を代表するキーワードを的確に設定して検索を行ったつもりでも、実際に表示される検索結果が多過ぎて、目的とする情報を取得するのに手間取るといったケースが多い。

【0004】このような問題を解消するため、目的語や述語からなる簡単なセンテンスを利用して情報検索を行うシステムが、例えば、特開平5-81326号公報、特開平5-108721号公報、特開平6-12451号公報等として既に提案されている。しかし、これらのものは、基本的に、センテンスを構成する構成要素の全てをアンド条件として検索を行う構成であるため、検索結果が思わしくないような場合には、改めてセンテンスを構成し直して最初から検索作業を実行しなければならず、特に、インターネットのように情報が過多である環境下では、検索が完了するまでの所要時間が非常に長くなってしまふといった問題が生じる可能性がある。

【0005】

【発明の目的】そこで、本発明の目的は、前記従来技術の欠点を解消し、インターネットの初心者であっても、必要とする情報のありかを短時間のうちに容易に知ることのできる情報検索システムおよび情報検索プログラムを提供することにある。

【0006】

【課題を解決するための手段】本発明の情報検索システムは、情報のありかとキーワードとを対応させて記憶したインデックスを利用し、入力されたキーワードに対応する情報のありかを検索して提示する情報検索システムであり、前記目的を達成するため、特に、センテンスとして入力されたキーワードを構文解析して主要な構成要素と他の構成要素とに分割するキーワード分析手段と、前記主要な構成要素のみをキーワードとして前記インデックスを検索することにより検索対象の母集合となる検索結果を求めて検索結果一覧テーブルに一時記憶させると共に、前記他の構成要素および該他の構成要素に代えて利用することのできる幾つかの置換可能要素をキーワードとして前記検索結果一覧テーブルを検索して前記検索結果一覧テーブル内の検索結果をグループ化する検索手段と、前記他の構成要素および前記幾つかの置換可能要素の中から択一的に要素を選択するための入力手段と、前記入力手段によって選択された要素に対応してグループ化されている検索結果の一覧を提示する検索結果確定手段とを備えたことを特徴とする構成を有する。

【0007】このような構成によれば、センテンスとして入力されたキーワードがキーワード分析手段によって構文解析されて主要な構成要素と他の構成要素とに分割され、まず、最初の段階で、主要な構成要素のみをキーワードとして検索手段による検索が行われ、その検索結果が母集合として検索結果一覧テーブルに一時記憶される。その後、検索結果を絞り込むために、主要な構成要素を除く他の構成要素および該他の構成要素に代えて利用することのできる幾つかの置換可能要素を用いた検索が前記検索手段によって行われ、その検索結果が、これらの他の構成要素および幾つかの置換可能要素に対応してグループ化される。この検索作業は、その検索範囲が検索結果一覧テーブルに限定されるため、短時間で行うことが可能である。ユーザは、この情報検索システムに設けられた入力手段を用いて、前記他の構成要素および前記幾つかの置換可能要素の中から択一的に要素を選択し、当該システムの検索結果確定手段に、選択した要素を認識させる。これを受けた検索結果確定手段は、選択された要素に対応してグループ化されている検索結果の一覧をユーザに提示する。このように、センテンスからなるキーワードの主要な構成要素を利用して所望する情報が存在する可能性のある範囲を母集合として特定し、更に、キーワードに含まれる他の構成要素、例えば、述語や修飾語等を利用した絞り込みのための検索が行われ

るので、インターネットの初心者であっても、必要とする情報のありかを短時間のうちに容易に知ることができる。また、絞り込みによって適切な情報が得られなかった場合には、再び前記他の構成要素あるいは前記幾つかの置換可能要素の中から別の要素を選択して絞り込みを行うことによって新たな情報の提示を求めることになるが、この段階では既に前記他の構成要素および前記幾つかの置換可能要素に対応する検索結果のグループ化が完了しているので、検索結果確定手段は、単に、ユーザが選択した要素に対応するグループの検索結果を検索結果一覧テーブルから読み込んで提示するだけでよく、実質的な検索処理は不要である。従って、特に、適切な情報が得られずに何回も絞り込みの操作を繰り返し実行するような状況下において、絞り込みの都度インデックス全体を最初から再検索する従来の情報検索システムに比べ、極めて高速の処理が可能である。更に、絞り込みの操作に適したキーワードが置換可能要素としてシステムの側から自動的に提供されるので、検索作業に不慣れなユーザであっても検索システムを容易に使いこなすことができる。

【0008】前述した置換可能要素は、前記検索手段に併設された置換可能要素記憶手段内に前記他の構成要素と対応させて記憶させることが可能である。

【0009】この場合、前記検索手段は、キーワード分析手段によって分割された他の構成要素毎に置換可能要素記憶手段を参照し、前記構成要素に代えて利用することのできる幾つかの置換可能要素を求めて検索結果一覧テーブルを検索し、前記検索結果一覧テーブル内の検索結果をグループ化する。

【0010】また、前記キーワード分析手段には、名詞および固有名詞を主要な構成要素として認識する主要構成要素認識機能を設けることができる。

【0011】一般に、センテンスとして入力されたキーワードに含まれる構成要素のうち、名詞および固有名詞が最も端的に目的とする情報の属性を表すので、これらのものをセンテンス中の主要な構成要素として認識するのが妥当である。

【0012】更に、検索結果確定手段には、検索手段の作動終了時に前記他の構成要素を自動的に選択する最適検索結果提示機能を備えることが望ましい。

【0013】このような構成を適用すれば、グループ化された複数の検索結果のうち、まず最初に、ユーザが指定したセンテンスに含まれる構成要素に対応した検索結果が提示されるので、高い確率で所望する情報のありかを引き当てることが可能となる。

【0014】また、前記検索手段には、分割された前記他の構成要素毎に検索優先度を判定し、検索優先度が高い前記他の構成要素およびこの構成要素に代えて利用することのできる幾つかの置換可能要素のみをキーワードとして前記検索結果一覧テーブルを検索して検索結果一

覧テーブル内の検索結果をグループ化する検索条件最適化機能を設けることができる。

【0015】このような構成を適用することにより、センテンスとして入力されたキーワードのうち優先度の高い構成要素とこの構成要素に関連する置換可能要素のみをキーワードとして検索およびグループ化の処理が行われるので、無駄な検索作業の実行が未然に防止され、高速な検索作業が達成される。

【0016】また、前記入力手段には、前記他の構成要素および前記幾つかの置換可能要素を並列してモニタ上に表示する選択可能要素一括表示機能を配備することができる。

【0017】グループ化による検索結果の絞り込みに利用されたキーワードの一覧がユーザに明示されるので、再検索のためのキーワードの設定に不慣れなユーザであっても検索システムを容易に使いこなすことができるようになる。

【0018】本発明の情報検索プログラムは、情報のありかとキーワードとを対応させて記憶したインデックスを利用し、入力されたキーワードに対応する情報のありかを検索して提示する情報検索システムの主要部を構成するコンピュータに設置される情報検索プログラムであり、前記目的を達成するため、特に、前記コンピュータにセンテンスとして入力されたキーワードを構文解析して主要な構成要素と他の構成要素とに分割するキーワード分析機能と、前記主要な構成要素のみをキーワードとして前記インデックスを検索することにより検索対象の母集合となる検索結果を求めて検索結果一覧テーブルに一時記憶させると共に、前記他の構成要素および該他の構成要素に代えて利用することのできる幾つかの置換可能要素をキーワードとして前記検索結果一覧テーブルを検索して前記検索結果一覧テーブル内の検索結果をグループ化する検索機能と、入力手段または情報伝達経路を介して入力される指令に応じて前記他の構成要素および前記幾つかの置換可能要素の内から択一的に要素を選択し、該選択した要素に対応してグループ化されている検索結果の一覧を出力する検索結果確定機能とを備えたことを特徴とする構成を有する。

【0019】このプログラムを設置することにより、通常のコンピュータを情報検索システムの主要部として利用することが可能となる。その機能に関しては、既に述べた通りである。

【0020】

【発明の実施の形態】以下、図面を参照して本発明の実施形態について詳細に説明する。図1は本発明を適用した一実施形態の情報検索システム1の主要部を簡略化して示した機能ブロック図である。

【0021】図1に示されるように、この情報検索システム1は、概略において、データ処理装置2とキーボード等の入力手段3、および、モニタあるいはプリンタ等

によって構成される出力装置4と、ハードディスクおよびRAM等からなる記憶装置5によって構成される。

【0022】このうち、データ処理装置2の部分は通常のパーソナルコンピュータあるいはワークステーション等のコンピュータによって構成されており、このコンピュータに設置（インストール）された情報検索プログラムとCPUとによって、キーワード分析手段6と検索手段7および検索結果確定手段8の機能が達成される。

【0023】図1の例では入力手段3をキーボード、また、出力装置4をモニタあるいはプリンタ等としているが、入出力装置は、これらのものに限定される必要はなく、データ処理装置2をウェブ上に設置して検索サーバとして利用するような場合においては、末端のユーザが操作する別のパーソナルコンピュータのキーボードやマウスがインターネット等の情報伝達経路を介して実質的な入力手段3として利用され、また、末端のユーザが操作する別のパーソナルコンピュータのモニタがインターネット等の情報伝達経路を介して実質的な出力装置4として利用されることになる。

【0024】記憶装置5には、データ処理装置2が構文解析に際して試用する辞書を格納した辞書記憶部9と、情報のありかをキーワードに対応させて記憶したインデックス10、および、検索結果を記憶するための検索結果一覧テーブル11（RAM内）と、各種の単語（述語、修飾語等）とこれに置き換えて利用することができる同種の単語とを対応させて記憶した置換可能要素記憶部12とが設けられている。

【0025】インデックス10には、データ処理装置2あるいは他の情報収集用コンピュータが通常の処理の空き時間を利用してウェブ上で実行するロボット検索等の情報収集処理によって、莫大な数のデータ、たとえば、ウェブ上のURLの値等がキーワードに関連付けて記憶されている。

【0026】図7は、置換可能要素記憶部12に記憶された置き換え可能な単語（置換可能要素）の一例を示す概念図であり、この例では、例えば、述語の一種である「見る」、「する」、「買う」が相互に置き換え可能な単語として定義され、また、修飾語の一種である「今日」、「明日」、「あさって」が相互に置き換え可能な単語として定義されている。

【0027】置き換え可能な単語の定義は、日本語の文法を満たすように設定してもよいし、また、その他の事情を考慮して必ずしも日本語の文法にはとらわれずに設定しても構わない。

【0028】次に、データ処理装置2のCPUが実行する情報検索処理（情報検索プログラム）の内容を示した図2～図5のフローチャートを参照してキーワード分析手段6、検索手段7、検索結果確定手段8等の機能について詳細に説明する。

【0029】まず、センテンスからなるキーワードを入

力してユーザがデータ処理装置2に検索処理を命じると、キーワード分析手段6としてのCPUは、この入力操作をステップs1の判定処理によって検出する。

【0030】この実施形態では、一例として、「今日、サッカーを見る」というキーワードが入力された場合を例にとって説明するものとする。

【0031】キーワードの入力を検知したCPUは、次いで、辞書記憶部9を参照して、このセンテンスを構文解析し、構成要素の単位にまで分割する（ステップs2）。構文解析自体については既に公知であるので説明を省略する。

【0032】この構文解析の結果、「今日、サッカーを見る」のセンテンスは、図8のs2、s3に示されるように、「今日／、／サッカー／を／見る」という5つの構成要素に分割されることになる。

【0033】次いで、キーワード分析手段6としてのCPUは、これらの構成要素の配列状態をテンプレートに記憶し（ステップs3）、更に、予め設定された検索優先度に従って、これらの構成要素の並べ替えを実行する（ステップs4）。

【0034】図6は本実施形態において適用した検索優先度について説明した概念図である。図6に示されるように、この実施形態では、「名詞と固有名詞」、「述語」、「修飾語」、・・・、「助詞」、・・・、「句読点」といった順で検索優先度が決められており、最も検索優先度の高い「名詞と固有名詞」が自動的に主要な構成要素として認識され（主要構成要素認識機能）、また、「名詞と固有名詞」を除いたものが自動的に他の構成要素として認識される。また、このうち、「助詞」と「助詞」よりも検索優先度の低いものが実質的に検索の必要のない他の構成要素として認識されるようになっていく。以上は一例であって、例えば、修飾語を更に形容詞と副詞とに分類して検索優先度を設定するといったことも技術的には可能である。

【0035】この実施形態においては、図6の検索優先度に従って、「今日／、／サッカー／を／見る」の配列が、図8のs4に示されるように「サッカー／見る／今日／を／、」の順に並べ替えられることになる。

【0036】次いで、CPUは、分割された構成要素の数を構成要素数記憶レジスタnに記憶し（ステップs5）、構成要素選択指標iに初期値1を設定する（ステップs6）。この場合、図8のs4に示されるようにnの値は5である。

【0037】そして、CPUは、構成要素選択指標iの現在値に基き、並べ替え実行後の配列から第i番目の構成要素を読み込み（ステップs7）、この構成要素が検索対象外のものであるか否か、要するに、「助詞」または「助詞」よりも検索優先度の低いものであるか否かを判定する（ステップs8）。

【0038】この段階ではi=1であり、最初に読み込

まれるのは「名詞」からなる主要な構成要素「サッカー」であるから、ステップs8の判定結果は偽となる。

【0039】よって、CPUは、構成要素選択指標iの値を1インクリメントし（ステップs9）、該指標iの現在値が構成要素の数nを越えているか否かを判定する（ステップs10）。

【0040】この例では、i=2、n=5であるからステップs10の判定結果は偽となり、CPUは、再び該指標iの現在値に基き、前記と同様にして、並べ替え実行後の配列から第i番目の構成要素を読み込む（ステップs7）。

【0041】ここで読み込まれるのは「述語」からなる「見る」であり、その検索優先度は「助詞」よりも高い。従って、その後の処理結果に関しては前記と同様である。

【0042】このようにして、順次インクリメントされる構成要素選択指標iの値に基いて、更に、「修飾語」からなる「今日」と、「助詞」からなる「を」がステップs7の処理で繰り返して読み込まれ、その時点における構成要素選択指標iの値は、図8のs6～s10で示されるとおり4となる。

【0043】「助詞」である「を」が読み込まれた段階でステップs8の判定結果は真となり、CPUは、構成要素選択指標iの値を1デクリメントして、この値を構成要素数記憶レジスタnに更新して記憶する（ステップs11）。図8のs6～s10で示されるとおり、構成要素選択指標iの現在値は4であるから、nの値は、図8のs11に示されるとおり3となる。

【0044】以上に述べたようにしてステップs7～ステップs10の処理を繰り返し実行することにより、実際に検索の対象とする必要のある構成要素が、並べ替え実行後の配列の先頭から何番目までに位置しているかが求められる。これが、本実施形態における検索条件最適化機能の主要部である。図8のs6～s10およびs11に示されるように、「助詞」または「助詞」よりも検索優先度の低い構成要素が検出されたときの構成要素選択指標iの値から1を減じた値が、実際に検索の対象とする必要のある構成要素の数であり、この例では、その値はn=3である。

【0045】なお、このような処理を繰り返す間にステップs10の判定結果が真となった場合、つまり、全ての構成要素が検索の必要のある対象として判定された場合には、最初にステップs5の処理で設定されたnの値には何らの変更操作も加えられないので、構成要素の数それ自体が検索の必要の有る要素数として認識されることになる。

【0046】次いで、CPUは、構成要素選択指標iの値を再び1に初期化し（ステップs12）、構成要素選択指標iの現在値に基いて、並べ替え実行後の配列から第i番目の構成要素を読み込み（ステップs13）、構成要素選択指標iの現在値が1であるか否か、つまり、

現時点で読み込まれている構成要素が最も検索優先度の高い「名詞または固有名詞」、要するに、主要な構成要素であるか否かを判定する（ステップs14）。

【0047】この段階では $i=1$ であり、読み込まれるのは「名詞」からなる「サッカー」であるから、主要構成要素認識機能実現手段としてのCPUは、「サッカー」を主要な構成要素として認識し、更に、検索手段7としてのCPUが、主要な構成要素である「サッカー」のみをキーワードとしてインデックス10の全領域を検索する（ステップs15）。

【0048】次いで、検索手段7としてのCPUは、「サッカー」をキーワードとする情報の有無がインデックス10に登録されているか否かを判定し（ステップs16）、検索結果が存在しない場合には、出力装置4にエラーを出力して検索が失敗したことをユーザに知らせ（ステップs36）、この情報検索処理を終了する。

【0049】一方、主要な構成要素である「サッカー」をキーワードとする情報の有無がインデックス10に登録されている場合には、検索手段7としてのCPUは、その検索結果全体を検索結果一覧テーブル11に転送し、検索対象の母集合として一時記憶させる（ステップs17）。これ以降の検索作業は検索結果一覧テーブル11（RAM）に記憶された母集合に対してのみ行われることになるので、インデックス10全体を検索する場合と比べて遥かに高速である。

【0050】次いで、CPUは構成要素数記憶レジスタ n の値が1であるか否かを判定する（ステップs18）。

【0051】ステップs18の判定結果が真となった場合には、最初にセンテンス入力されたキーワードが「名詞または固有名詞」からなる主要な構成要素のみによって構成されているか、あるいは、主要な構成要素を除く他の構成要素の全ての検索優先度が「助詞」よりも低く、主要な構成要素を除く他の構成要素に関連する検索処理が不要であることを意味するので、これ以上の検索処理を行っても意味が無い。

【0052】従って、この場合、検索結果確定手段8としてのCPUは、検索結果一覧テーブル11に一時記憶された全ての検索結果を出力装置4に出力してユーザに提示し（ステップs37）、この情報検索処理を終了する。

【0053】また、ステップs18の判定結果が偽となった場合には、主要な構成要素を除く他の構成要素が存在し、しかも、そのうち1つ以上のものが検索の対象とする必要のある構成要素であることを意味するので、CPUは、更に、構成要素選択指標 i の値を1インクリメントし（ステップs19）、該指標 i の現在値に基いて、前記と同様にして、並べ替え実行後の配列から第 i 番目の構成要素を読み込む（ステップs13）。

【0054】 $i=2$ の現時点で読み込まれる構成要素は

「述語」からなる「見る」であるから、ステップs14の判定結果は必然的に偽となる。

【0055】従って、検索手段7としてのCPUは、検索を必要とする他の構成要素の1つである「見る」のみをキーワードとして、ステップs17の処理で記憶された全体の検索結果、つまり、検索結果一覧テーブル11（RAM）の母集合のみを検索する（ステップs20）。

【0056】次いで、CPUは、「見る」をキーワードとする情報の有無が検索結果一覧テーブル11に記憶されているか否かを判定し（ステップs21）、検索結果が存在すれば、検索条件最適化機能実現手段としてのCPUは、並べ替え実行後の配列における第 i 番目の構成要素に置き換えて利用することができる同種の単語（置換可能要素）を図7の置換可能要素記憶部12から求め、その単語を第 i 番目の構成要素と共に記憶する（ステップs22）。

【0057】ここでは構成要素選択指標 $i=2$ の値に応じて「見る」が選択されているので、図7の置換可能要素記憶部12を参照して、「する」という単語と「買う」という単語が一時記憶されることになる。

【0058】次いで、CPUは、構成要素選択指標 i の値を1インクリメントし（ステップs23）、該指標 i の値が構成要素数記憶レジスタ n の値を超えているか、つまり、検索の対象となる他の構成要素が未だ残っているか否かを判定する（ステップs24）。

【0059】この段階では構成要素選択指標 i の現在値は3であり、構成要素数記憶レジスタ n の値である3を超えてはいないので、CPUは、構成要素選択指標 i の値に基いて、再び、並べ替え実行後の配列から $i=3$ 番目の構成要素、つまり、「修飾語」である「今日」を読み込み（ステップs13）、ステップs14の判定処理を実行した後、検索を必要とする他の構成要素の1つである「今日」のみをキーワードとして、検索結果一覧テーブル11（RAM）の母集合のみを検索する（ステップs20）。

【0060】次いで、CPUは、「今日」をキーワードとする情報の有無が検索結果一覧テーブル11に記憶されているか否かを判定し（ステップs21）、検索結果が存在すれば、検索条件最適化機能実現手段としてのCPUは、並べ替え実行後の配列における第 i 番目の構成要素に置き換えて利用することができる同種の単語（置換可能要素）を図7の置換可能要素記憶部12から求め、その単語を第 i 番目の構成要素と共に記憶する（ステップs22）。

【0061】ここでは構成要素選択指標 $i=3$ の値に応じて「今日」が選択されているので、図7の置換可能要素記憶部12を参照して、「明日」という単語と「あさって」という単語が一時記憶されることになる。

【0062】次いで、CPUは、構成要素選択指標 i の

値を1インクリメントし(ステップs23)、該指標iの値が構成要素数記憶レジスタnの値を超えているか、要するに、検索の対象となる他の構成要素が未だ残っているか否かを判定する(ステップs24)。

【0063】この段階では構成要素選択指標iの現在値は4であり、構成要素数記憶レジスタnの値である3を超えているので、CPUは、構成要素選択指標iに構成要素数記憶レジスタnの値を再設定して、構成要素選択指標iの値を検索結果が存在する構成要素の数、言い換えれば、検索結果の表示が可能な構成要素の数に合わせる(ステップs26)。

【0064】一方、このような処理を繰り返し実行する間にステップs21の判定結果が偽となった場合には、並べ替え実行後の配列の第i番目から読み込まれた構成要素に対応する検索結果が存在しないことを意味する。この場合、CPUは、指標iの現在値から1を減じ、その値を構成要素選択指標iに再設定して、構成要素選択指標iの値を検索結果が存在する構成要素の数に合わせる(ステップs25)。仮に、第3番目の構成要素である「今日」に対応する検索結果が存在しなかったとすれば、 $3-1=2$ であり、検索結果が存在する構成要素は2つのみとなる。

【0065】次いで、検索手段7としてのCPUは、並べ替え実行後の配列における第2番目の構成要素およびこれに対応して記憶された置換可能な同種の単語と、並べ替え実行後の配列における第3番目の構成要素およびこれに対応して記憶された置換可能な同種の単語と、・・・、並べ替え実行後の配列における第i番目の構成要素およびこれに対応して記憶された置換可能な同種の単語との組み合わせによって生成される全てのキーワードを用いて検索結果一覧テーブル11(RAM)に記憶された母集合に対して検索処理を実行し、キーワード毎に検索結果をグループ化する(ステップs27)。

【0066】従って、この例では、第2番目の構成要素およびこれに対応して記憶された置換可能な同種の単語である「見る」、「する」、「買う」と、第i=3番目の構成要素およびこれに対応して記憶された置換可能な同種の単語である「今日」、「明日」、「あさって」との組み合わせにより、図9および図10に示されるように $3 \times 3 = 9$ 種のキーワードが生成され、検索結果一覧テーブル11(RAM)の内容がキーワード毎に9種のグループに分けられることになる。但し、検索結果一覧テーブル11の内容は既に「サッカー」を含むものに限定されているので、ここでいうキーワードには、全てのキーワードに共通する並べ替え実行後の配列上における第1番目の構成要素(主要な構成要素)である「サッカー」は含まれない。

【0067】次いで、CPUは、並べ替え実行後の配列上における第2番目から第i番目に至る構成要素を選択可能な状態として、各構成要素の各々を、ステップs3

で記憶したテンプレートに沿って、並べ替え実行前の位置でモニタ等の出力装置4に表示し(ステップs28)、まず最初に、最適検索結果提示機能実現手段としてのCPUが、並べ替え実行後の配列における第2番目の構成要素と、並べ替え実行後の配列における第3番目の構成要素と、・・・、第i番目の構成要素とによって生成されたキーワード、つまり、最初にユーザがセンテンス入力したキーワードに相当するグループの検索結果を出力装置4に表示する(ステップs29)。

【0068】図11は出力装置4上での表示状態の一例を示した概念図である。このように、出力装置4上には、前述した検索処理の要不要とは関係なく、構成要素の全てが、ステップs3で記憶したテンプレートに沿って「今日／、／サッカー／を／見る」の順序で表示され、このうち、並べ替え実行後の配列における第2番目～第i=3番目(但し、iの値は最終値)までの構成要素である「見る」と「今日」の部分が他の単語と選択的に置き換え可能とされる。

【0069】ここで、仮に、「今日」に対応する検索結果が存在していなかったとするなら、iの値は2となるから、「見る」の構成要素のみが選択可能な状態で表示されることになる。

【0070】次いで、CPUは、選択可能な構成要素に対してユーザによるマウスクリック操作等が行われるか(ステップs30)、あるいは、ユーザからの検索終了指令が入力されるのを待つ待機状態に入る(ステップs35)。

【0071】そして、選択可能な構成要素に対して入力手段3を介してマウスクリック操作等が行われた場合には(ステップs30)、選択可能要素一括表示機能実現手段としてのCPUが、選択された構成要素に対応する置き換え可能な単語(置換可能要素)の情報を内包したポップアップメニューを表示して(ステップs31)、ユーザからの選択操作を待機する(ステップs32)。

【0072】図12は、図11に示されるような出力装置4の表示状態において、ユーザが構成要素「見る」をクリックした場合の画面変化の一例を示した概念図である。この場合、「見る」に対応して記憶された置き換え可能な単語(置換可能要素)である「する」と「買う」がポップアップメニュー上に並列して一括表示され、選択可能な要素としてユーザに提示されることになる。

【0073】ついで、ユーザが入力手段3を介してマウスクリック操作等を実行し、置き換え可能な単語の何れかを選択すると(ステップs32)、CPUは、新たに選択された置き換え可能な単語(置換可能要素)を当該ポップアップメニューの基部位置に書き替えて表示し(ステップs33)、検索結果確定手段8としてのCPUが、現時点で出力装置4に表示されている要素の組み合わせキーワードに対応してグループ化されている検索結果の内容を出力装置4に表示してユーザに提示する

(ステップ s 34)。

【0074】図13は、図12に示されるような出力装置4の表示状態において、ユーザが要素「する」をクリックした場合の画面変化の一例を示した概念図である。この場合、構成要素「見る」に代えて置き換え可能な単語(置換可能要素)「する」が選択されることになるので、それまでユーザに提示されていた「サッカー／見る／今日」(今日、サッカーを見る)に対応するグループの検索結果に代えて、「サッカー／する／今日」(今日、サッカーをする)に対応するグループの検索結果が更新して表示されることになる。

【0075】その後、CPUは、再び、選択可能な構成要素に対するマウスクリック操作、あるいは、検索終了指令の入力を待つ待機状態に戻る(ステップ s 30、ステップ s 35)。

【0076】そして、ユーザが入力手段3を介して再びマウスクリック操作等を実行すると、CPUは、前記と同様の処理を繰り返し実行し、改めて生成された要素の組み合わせキーワードに対応してグループ分けされている検索結果の内容を出力装置4に表示してユーザに提示する。

【0077】このように、要素の組み合わせによるキーワードの変更は何度でも可能であり、その都度、このキーワードに対応するグループの検索結果が表示されることになるが、検索結果のグループ分けに関連した処理は既にステップ s 27の時点で完了しており、ステップ s 30～ステップ s 35に至る再検索処理では、単に、グループ化された検索結果の内容を表示するだけでよいので、再検索および再表示に関連する処理は極めて短時間で行うことができる。

【0078】しかも、絞り込みの検索に利用するためのキーワード(要素)はポップアップメニューによって一括してユーザに提示されるので、検索結果の絞り込みに使用するキーワードの設定に不慣れなユーザであっても、絞り込みを利用した検索を容易に実行することができる。

【0079】最終的に、ユーザが入力手段3を介して検索終了指令を入力した段階で(ステップ s 35)、この情報検索処理は終了する。

【0080】

【発明の効果】本発明の情報検索システムおよび情報検索プログラムは、センテンスとして入力されたキーワードをキーワード分析手段で構文解析することによって主要な構成要素と他の構成要素とに分割し、まず、最初の段階で主要な構成要素のみをキーワードとして検索手段による検索を実行して検索対象とする母集合を求めて検索結果一覧テーブルに一時記憶し、その後、検索結果を絞り込むためのキーワードの一部を構成する他の構成要素およびこれに代替可能な置換可能要素を用いて母集合の範囲内で検索を実行するようにしているので、検索結

果の絞り込みを行う際に大容量のインデックスが再検索されることはなく、詳細な検索条件を利用したキーワード検索を短時間のうちに実施することができる。また、検索結果を絞り込むためのキーワードとなる他の構成要素およびこれに代替可能な置換可能要素に関連する検索結果は検索処理によって予めグループ化され、その後、他の構成要素および置換可能要素の組み合わせを利用してキーワードでグループを選択することによって自由に検索結果を読み出すことができるので、重複した再検索作業を何度も行う必要がなく、所望するキーワードに対応する情報のありかを短時間のうちに知ることができる。しかも、絞り込みの操作に適したキーワードが置換可能要素としてシステムの側から自動的に提供されるので、検索作業に不慣れなユーザであっても検索システムを容易に使いこなすことができる。

【0081】更に、検索手段の作動終了時にはユーザがセンテンス入力したキーワードと同等の構成要素の組み合わせに対応した検索結果が最初に提示されるので、高い確率で所望する情報のありかを引き当てることが可能である。

【0082】また、分割された他の構成要素毎に検索優先度を判定し、検索優先度が高い他の構成要素およびこれに対応する置換可能要素のみをキーワードとして検索結果をグループ化するようにしているので、絞り込み検索に効果の少ない無意味な単語によるグループ化処理が未然に禁止され、グループ化処理を含む全体の検索作業の所要時間の短縮が実現される。

【0083】更に、グループ化による検索結果の絞り込みに利用されたキーワードの一覧、つまり、他の構成要素と幾つかの置換可能要素の一覧がユーザに明示される構成であるため、再検索のためのキーワードの設定に不慣れなユーザであっても、提示された単語を選択するだけの簡単な操作で検索システムを容易に使いこなすことができる。

【図面の簡単な説明】

【図1】本発明を適用した一実施形態の情報検索システムの主要部を簡略化して示した機能ブロック図である。

【図2】データ処理装置のCPUが実行する情報検索処理の内容を示したフローチャートである。

【図3】データ処理装置のCPUが実行する情報検索処理の内容を示したフローチャートの続きである。

【図4】データ処理装置のCPUが実行する情報検索処理の内容を示したフローチャートの続きである。

【図5】データ処理装置のCPUが実行する情報検索処理の内容を示したフローチャートの続きである。

【図6】検索優先度の一例について説明した概念図である。

【図7】置換可能要素記憶部に記憶された置き換え可能な単語の一例を示した概念図である。

【図8】構文解析の一例について示した概念図である。

【図9】グループ分けのキーワードの一例を示した概念図である。

【図10】グループ分けのキーワードと情報の有りととの対応関係の一例を示した概念図である。

【図11】出力装置上での表示状態の一例を示した概念図である。

【図12】ユーザが構成要素をクリックした場合の画面変化の一例を示した概念図である。

【図13】ユーザが構成要素をクリックして検索条件を変更した場合の画面変化の一例を示した概念図である。

【符号の説明】

1 情報検索システム

2 データ処理装置（コンピュータ）

3 入力手段

4 出力装置

5 記憶装置

6 キーワード分析手段

7 検索手段

8 検索結果確定手段

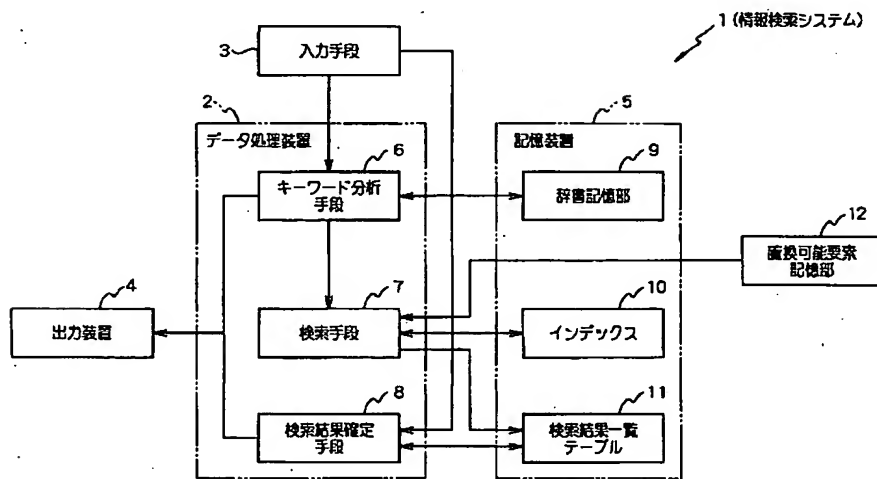
9 辞書記憶部

10 インデックス

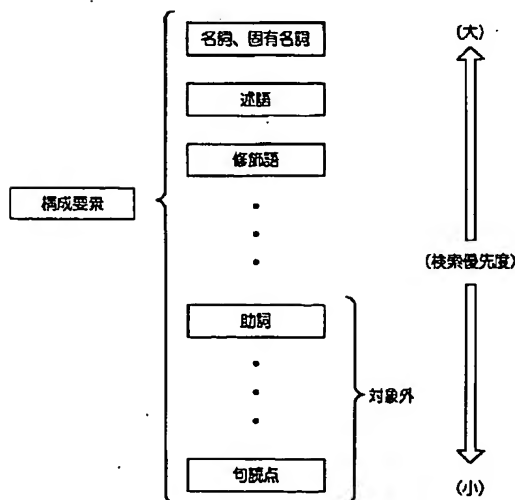
11 検索結果一覧テーブル

12 置換可能要素記憶部

【図1】



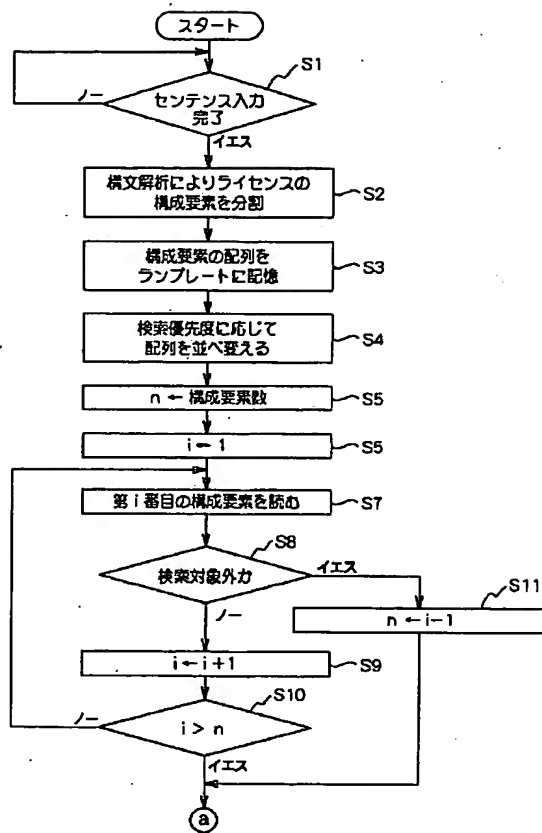
【図6】



【図7】

構成要素	置換可能要素
見る	する、買う、・・・
する	見る、買う、・・・
買う	見る、する、・・・
...	...
今日	明日、あさって、・・・
明日	今日、あさって、・・・
あさって	今日、明日、・・・
...	...

【図2】



【図8】

S1 今日、サッカーを見る

S2, S3 今日 /、 / サッカー / を / 見る

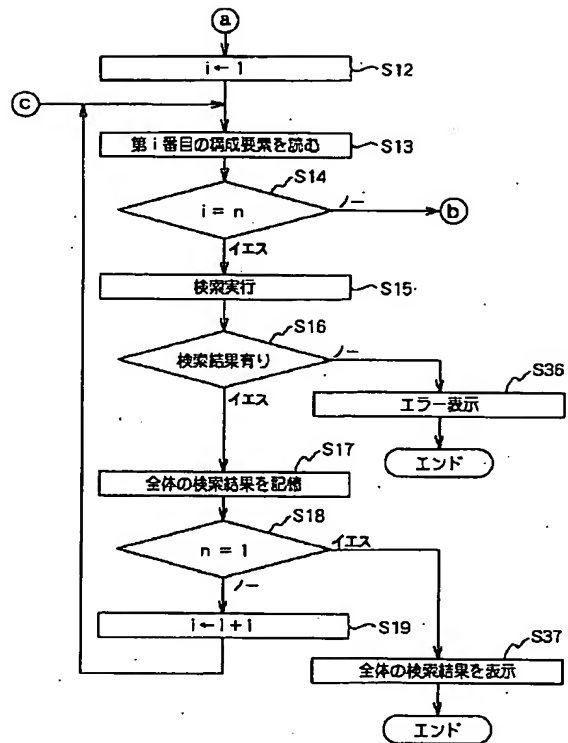
S4 サッカー / 見る / 今日 / を /、

S5 $n (=5)$

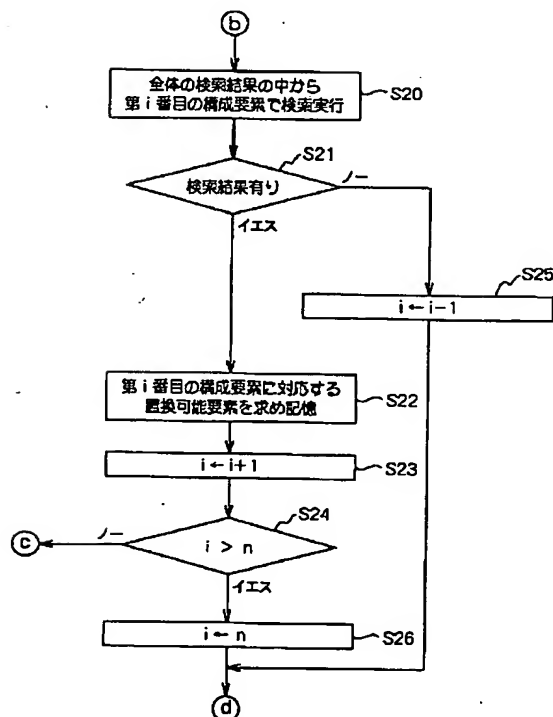
S6~S10 サッカー / 見る / 今日 / を /、
 $i = 4$

S11 サッカー / 見る / 今日 / を /、
 $n = i - 1 = 3$

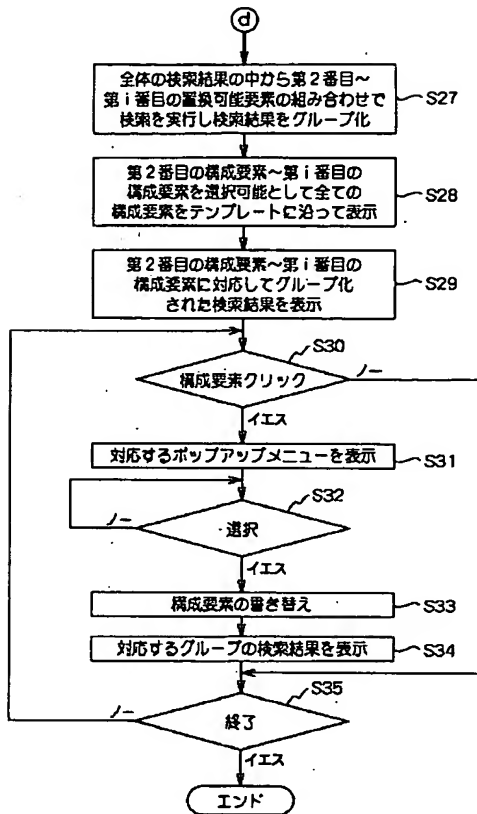
【図3】



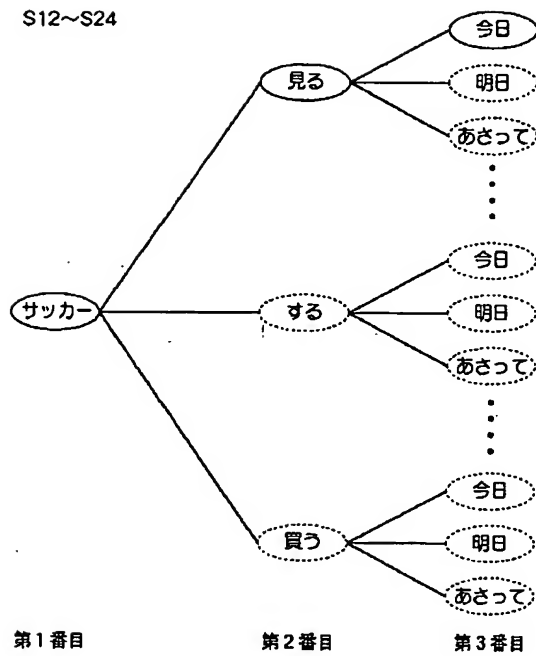
【図4】



【図5】



【図9】



【図10】

組み合わせ	URL
サッカー/見る/今日 (今日、サッカーを見る)	URL ₁₁ , URL ₁₂ , URL ₁₃ , . . .
サッカー/見る/明日 (明日、サッカーを見る)	URL ₂₁ , URL ₂₂ , URL ₂₃ , . . .
サッカー/見る/あさって (あさって、サッカーを見る)	URL ₃₁ , URL ₃₂ , URL ₃₃ , . . .
サッカー/する/今日 (今日、サッカーをする)	URL ₄₁ , URL ₄₂ , URL ₄₃ , . . .
サッカー/する/明日 (明日、サッカーをする)	URL ₅₁ , URL ₅₂ , URL ₅₃ , . . .
サッカー/する/あさって (あさって、サッカーをする)	URL ₆₁ , URL ₆₂ , URL ₆₃ , . . .
サッカー/買う/今日 (今日、サッカーを買う)	URL ₇₁ , URL ₇₂ , URL ₇₃ , . . .
サッカー/買う/明日 (明日、サッカーを買う)	URL ₈₁ , URL ₈₂ , URL ₈₃ , . . .
サッカー/買う/あさって (あさって、サッカーを買う)	URL ₉₁ , URL ₉₂ , URL ₉₃ , . . .
⋮	⋮

【図11】

今日、サッカーを見る

▼今日、 、 サッカー を ▼見る

検索結果

www.###.###.### 今日サッカーを見る場合は・・・

www.###.###.### 今日やっていたサッカーを見ると・・・

www.###.###.### 今日、サッカーを見るときに・・・

【図12】

今日、サッカーを見る

▼今日、 、 サッカー を ▼見る
する✓
買う

検索結果

www.###.###.### 今日サッカーを見る場合は・・・

www.###.###.### 今日やっていたサッカーを見ると・・・

www.###.###.### 今日、サッカーを見るときに・・・

【図13】

今日、サッカーを見る

▼今日、 、 サッカー を ▼する

検索結果

www.###.###.### 今日のサッカーの結果を表示すると・・・

www.###.###.### 今日もサッカーをするばかりで・・・

www.###.###.### 今日はサッカーをすると彼に・・・